PREE_Assignment

Andres Mancera Barreto

2023-09-19

Title

The islanders: a mock analysis of penguin distribution in three different islands.

Abstract

What can I say? There is available data on penguins in R that is not mine but is free to use. An assignment about data reproducibility requires you to copy what I have done. I am trying to accomplish that and not destroy my chances of graduating. I hope it works.

Introduction

You know how introductions go. Start broad on the topic, then narrow it down without pushing it too much. Insert claims from others that help support your thesis (Carey & Alexander, 2003; Cahill *et al.*, 2013; Harvey *et al.*, 2023). Then repeat it, make it more specific, and add more citations (Baker, 2006; Barlow *et al.*, 2020; Acuña-Tarazona *et al.*, 2022).

Lastly, finish it with a statement on the purpose of the paper and study, which in this case is to create a repository and reproducible research that you can, hopefully, repeat. Related to the data, we hypothesize that there is a higher probability of finding one penguin species on the island where we are surveying.

Methods

Using data in the *Palmer Penguins* package in R, we extracted the data to two separate CSV files and stored them in a *data* folder. The one named "penguins" was selected from these data frames to do some data manipulation, result production and storage.

Firstly, we created a frequency table that showed us the relative frequency of each species of penguin that was scouted and studied at the Palmer research station. This was done to create a plot and store the data in a tidy data folder.

The barplot for the proportion of penguins was created using the function *geom_bar* from the *ggplot* package.

We (I guess I am not used to the I) proceeded to do a contingency table for the penguins found per island and store it in the same *tidy data* folder. From this data, another barplot was created similarly to the one above.

The first part of this is redundant. I'm just not very smart, so yeah.... also, I'm too lazy to erase it, but not enough to write about it. Even though that would have been faster and easier?

Another contingency table was produced from the penguin's data to ensure that we did not violate assumptions for a χ^2 contingency test, and stored in the figures directorially, a χ^2 contingency test was done using the *chisq.test* function from the *janitor* package. That's it.
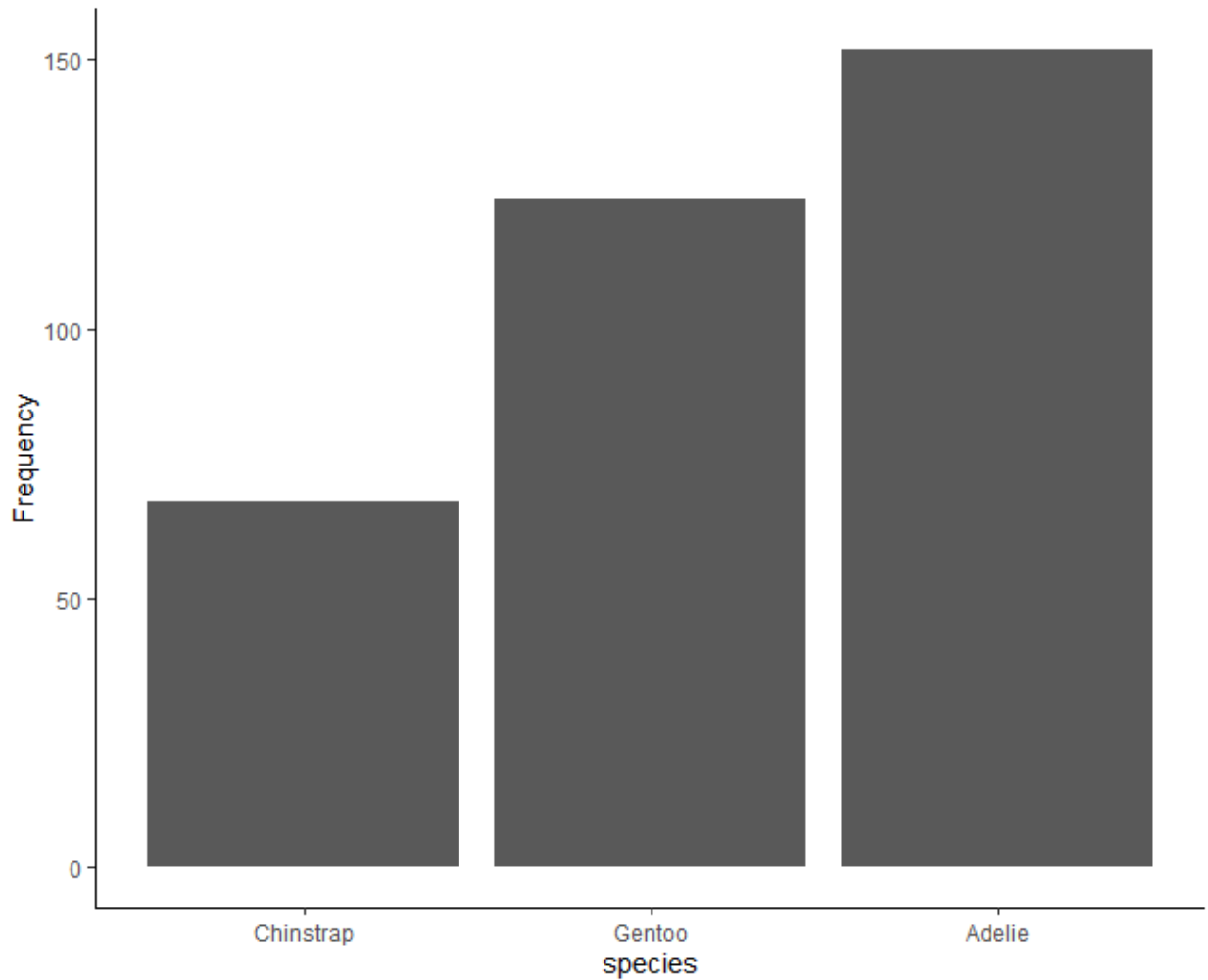
Results

Here are my results.

Figure 1: Species frequency in the penguin's dataset.

It appears that Adelie penguins are the most abundant in the dataset. Now let's see if there are differences among islands:

Table 1: Individuals count per island in the "palmerpenguins" package.

island	Adelie	Chinstrap	Gentoo	Total
Biscoe	44	0	124	168
Dream	56	68	0	124
Torgersen	52	0	0	52
Total	152	68	124	344

And in a graph:

We did not violate any assumptions for the χ^2 test, so we can say that the probability of finding one penguin species is significantly associated with the island where we are surveying (χ^2 contingency test; $df = 4$; $\chi^2 =$
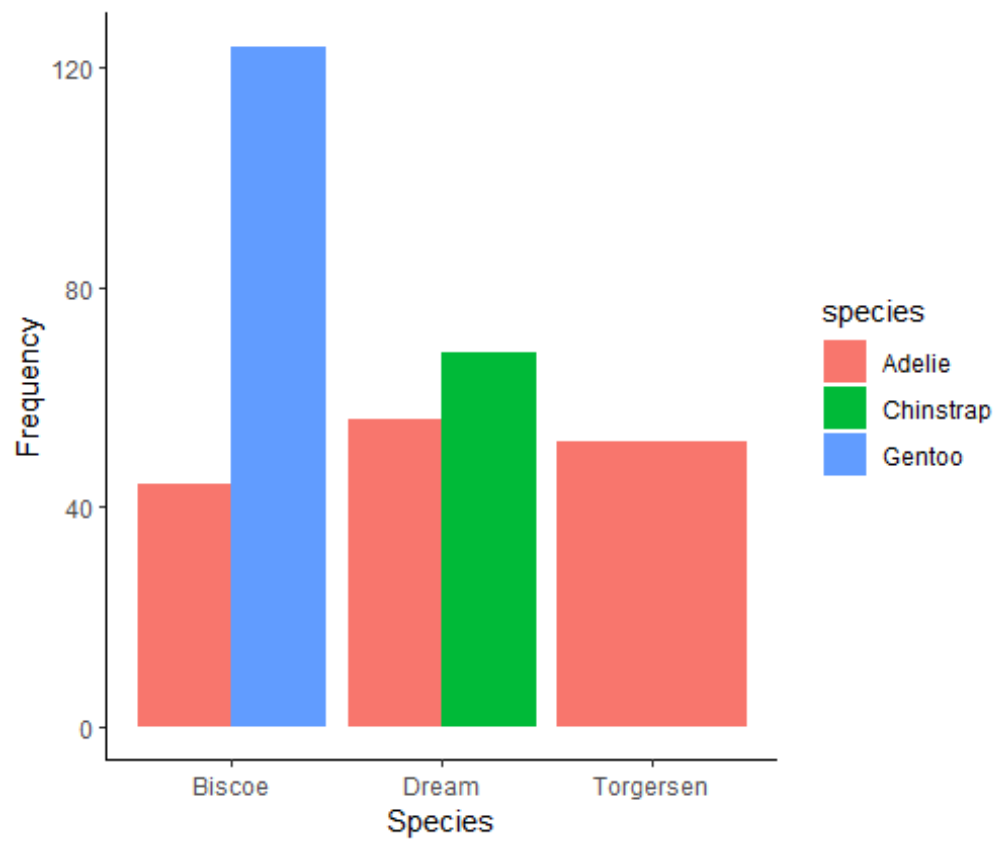
Figure 2: Species distribution across islands in the 'palmerpenguins' package.

299.55; $P < 0.001$). Based on our bar plot (Fig. 2), the distribution of penguins is vastly different among islands.

Cheers.

Discussion

We found that penguin species are related to the island where they are surveyed. Shocking. I do not know what else to say, mate. *Torgersen* is everywhere, and *Dream* and *Biscoe* are locked to particular islands. Neat.

As in any discussion, you also need to include some citation that supports your take and data (Anderson, 1984; Alvarez, 2002), so that there is a background and connection to similar studies (Boff *et al.*, 2015). In this case, as through the entire documents, all the references are non-related and only included as an exercise. None of the people here referenced have said anything remote similar to what I am quoting Acuña-Tarazona *et al.* (2022).

Then, you finish with a bombastic statement or suggestions for future studies. Probably something someone in your lab is already working on, creating a nice circular experience of knowledge in the lab.

Tables

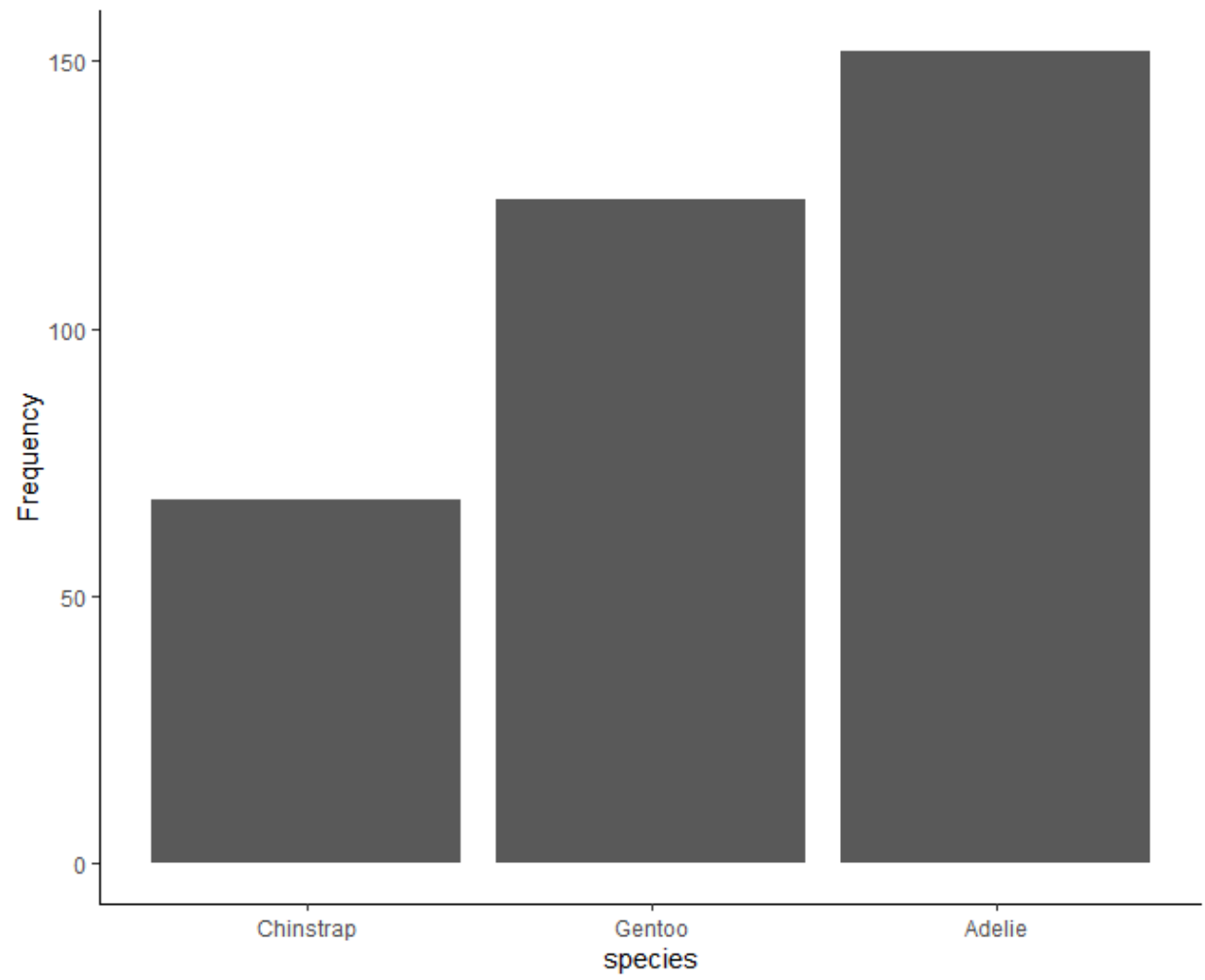Table 1: Individuals count per island in the "palmerpenguins" package.

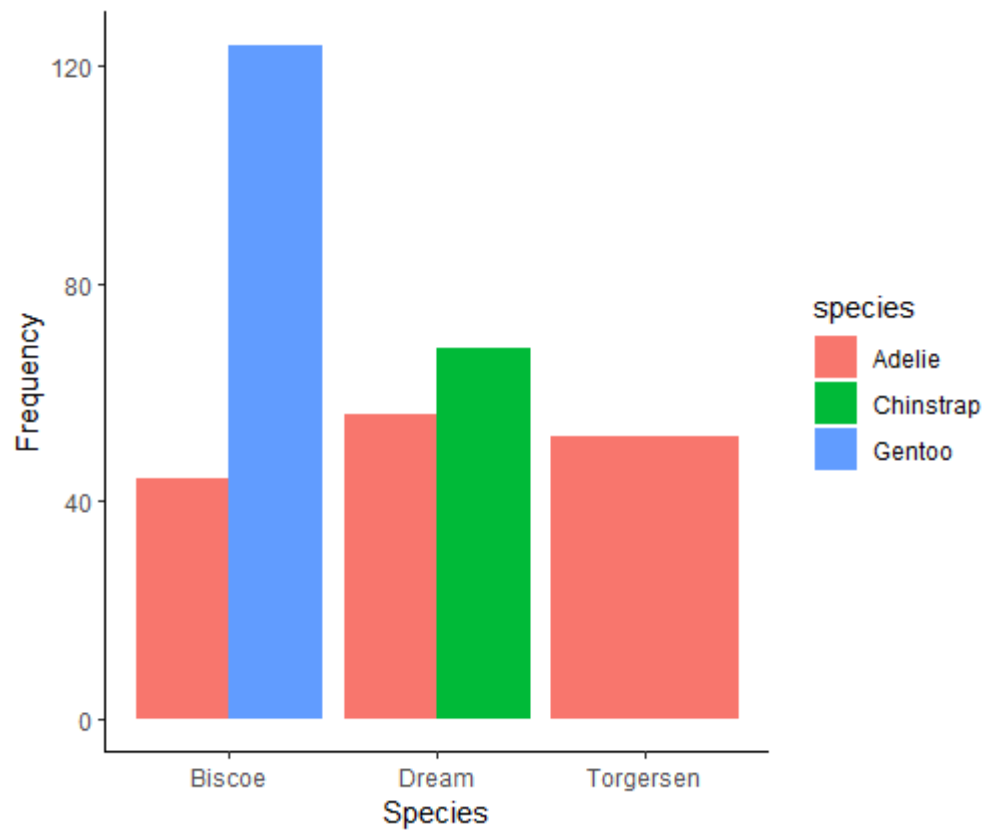| island	Adelie	Chinstrap	Gentoo	Total
Biscoe	44	0	124	168
Dream	56	68	0	124
Torgersen	52	0	0	52
Total	152	68	124	344

Figure captions

Fig. 1: Species frequency in the penguin's dataset.

Fig. 2: Species distribution across islands in the 'palmerpenguins' package.

Figures

Appendices

See the PREE.R script for how this was done.

References

We used R version 4.3.1 (R Core Team, 2023) and the following R packages: palmerpenguins v. 0.1.1 (Horst *et al.*, 2020), rmarkdown v. 2.24 (Xie *et al.*, 2018, 2020; Allaire *et al.*, 2023), tidyverse v. 2.0.0 (Wickham *et al.*, 2019).

Acuña-Tarazona, M., Mehltreter, K., Toledo-Aceves, T., Sosa, V.J., Flores-Palacios, A. & Kessler, M. (2022) Effects of microenvironmental factors on the diversity and composition of fern and orchid assemblages in an Andean paramo in Peru. *Flora*, **293**, 152107.

Allaire, J., Xie, Y., Dervieux, C., McPherson, J., Luraschi, J., Ushey, K., *et al.* (2023) *rmarkdown: Dynamic documents for r.*

Alvarez, M.D. (2002) Illicit crops and bird conservation priorities in Colombia. *Conservation Biology*, **16**, 1086–1096.

Anderson, M. (1984) The evolution of eusociality. *Annual review of ecology and systematics*, **15**, 165–189.

Baker, W.L. (2006) Fire and restoration of sagebrush ecosystems. *Wildlife Society Bulletin*, **34**, 177–185.

Barlow, N.L., Kirol, C.P., Doherty, K.E. & Fedy, B.C. (2020) Evaluation of the Umbrella Species Concept at Fine Spatial Scales. *Journal of Wildlife Management*, **84**, 237–248.

Boff, S., Forfert, N., Paxton, R.J., Montejo, E. & Quezada-Euan, J.J.G. (2015) A behavioral guard caste in a primitively eusocial orchid bee, Euglossa viridissima, helps defend the nest against resin theft by conspecifics. *Insectes Sociaux*, **62**, 247–249.

Cahill, A.E., Aiello-Lammens, M.E., Fisher-Reid, M.C., Hua, X., Karanewsky, C.J., Yeong Ryu, H., *et al.* (2013) How does climate change cause extinction? *Proceedings of the Royal Society B: Biological Sciences*, **280**, 20121890.

Carey, C. & Alexander, M.A. (2003) Climate change and amphibian declines: Is there a link? *Diversity and Distributions*, **9**, 111–121.

Harvey, J.A., Tougeron, K., Gols, R., Heinen, R., Abarca, M., Abram, P.K., *et al.* (2023) Scientists' warning on climate change and insects. *Ecological Monographs*, **93**.

Horst, A.M., Hill, A.P. & Gorman, K.B. (2020) *palmerpenguins: Palmer archipelago (antarctica) penguin data.*

R Core Team. (2023) *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., *et al.* (2019) Welcome to the tidyverse. *Journal of Open Source Software*, **4**, 1686.

Xie, Y., Allaire, J.J. & Grolemund, G. (2018) *R markdown: The definitive guide.* Chapman; Hall/CRC, Boca Raton, Florida.

Xie, Y., Dervieux, C. & Riederer, E. (2020) *R markdown cookbook.* Chapman; Hall/CRC, Boca Raton, Florida.